COLLABORATIVE GAME-BASED LEARNING: USING TEXT COLLECTING TECHNOLOGY TO ANALYZE LEARNERS' PERFORMANCE

By

Ruobin Wang

Senior Thesis in Computer Engineering

University of Illinois at Urbana-Champaign

Advisor: Christopher Schmitz, Michael Tissenbaum

May 2021

Abstract

In a collaborative learning environment, instructors need to pay attention to individual performance and experience throughout the activity. During the activity, instructors usually walk around the classroom to get feedback from each group of participants one by one orally; after the activity, feedback is usually in the form of surveys, quizzes, and assignments. The feedback provides an inconsistent understanding of participants ’behaviors and feelings due to limitations on their recall of all the details during the learning process and of the content of all the questions on the surveys or assignments. In this thesis, we describe a real-time, dynamic, minimalist feedback system designed for collaborative game-based learning environments using data-with-scale and keywords. This newly developed educational tool allows instructors to track the performances of each participant during the activity with a real-time analyzing viewing panel and produce a comprehensive final report. The instructors can view the topics being discussed in each group; receive notifications for questions, concerns, off-topic conversations; and see the main content of the dialogue between participants on an interactive platform similar to a control panel to better understand the learning curve and difficulties of the participants in such uncontrollable and varied situations.

Subject Keywords: Computer-supported collaborative learning, Collaborative game-based learning, Game-based learning environment, Dialogue analysis

Acknowledgments

I really appreciate the advice and help I received from Professor Christopher Schmitz and Professor Michael Tissenbaum when working on the research project. It has been a challenging but fruitful experience for the past year, especially during this period of Covid when communication becomes harder. I am grateful for all the support from my mentors and my colleagues in the City Settlers group.

Contents

1. Introduction 1
2. Literature Review 2
3. Problem Description and Limitations 4
   3.1 Bias on Game-based Learning 4
   3.2 Difficulties on Assessing Game-based Learning Environment 4
   3.3 Limitations 4
      3.3.1 Speech Separation 4
      3.3.2 Instant Feedback 5
4. Methods & Design 6
   4.1 Real-time Feedback 6
   4.2 Comprehensive Performance Analysis 7
      4.2.1 Topic Analysis 7
      4.2.2 Sentiment Analysis 7
      4.2.3 Turn of Talks 8
      4.2.4 Problem Solving 9
   4.3 Interactive Access to Data Collection 9
      4.3.1 Panel View 9
      4.3.2 Hotwords View 10
      4.3.3 Active Time View 11
5. Conclusion 11
References 13

1. Introduction

This research is designed to support the City Settlers project led by Professor Tissenbaum which is an educational game designed for middle school students to learn interdisciplinary ecological, historical and economical principles. The three main features are the collaborative and competitively built cities within groups of participants where they will set a goal together and learn how to work together through the run-in period; the independent ecosystem inside the game, namely the logic behind the game where every choice made by participants has consequences; and the connection between STEM and social contents, simulating real scenarios in the game.

2. Literature Review

Over 98 percent of all schools own advanced technologies such as computers, cable TVs, and CD-ROM driver [1]. Technologies are involved in teaching mathematics, languages, and history throughout kindergarten to college. With the help of these new technologies, teachers are able to implement new teaching formats such as games, online learning platforms, interactive whiteboards, class blogs and wikis. Among all the new teaching formats, educational gaming is one of the most challenging choices. Even though most teenagers love playing games, the integration in between needs to be comprehensively considered. It is hard for the instructors to control games that involve cooperative learning and highly interactive tasks. As a result, data collection for every individual is important for teachers to keep up with students 'performance. The following paragraphs summarize three published works that address important points to consider when filtering the data.

In the research report [2], the authors mention two significant concepts of higher education—result and quality—which should be measured continuously. In their ideal game design, they measured the student performance based on usage (how long and how often students participate in the game, mostly questionnaire form) and assessment after playing the game (quizzes or exams based on the knowledge taught through the game). These two methods can provide continuous data on the predetermined areas and individual feedback for the research team. To a higher level, the game should have an automated assessment system that makes sure students understand certain concepts before proceeding to the next step. So instead of tracking the performance of learners manually, the automated system can collect the detailed information on each task. The idea of "penalty" is interesting since penalty is rare in games, especially for young learners. However, more training to ensure learner progress in any learning scenario is necessary, and an appropriate amount of "penalty" prevents learners from guessing until they get the right answer.

According to the conference paper [3], one of the most significant questions of game-based collaborative learning is how to detect off-task behavior of the learners. The research mainly focuses on building off-task dialogue models and using them to identify chat messages that are not related to topics. Pre-trained word2vec embeddings are used to capture semantic and syntactic features of individual words. Then embeddings are derived from the ELMo and BERT models with lexical semantics collected. From the embeddings, every individual word of the learners and additional information about how the words are used in specific contexts will be recorded. The models are built based on the chat history of a collaborative game-based learning environment on ecosystem science from CRYSTAL

ISLAND: ECOJOURNEYS. Students are divided into groups with facilitators who ask leading questions when the learners are off-topic. The research team collects datasets from both the learners and facilitators. The extraction of the content has several special features: number of times that students have participated in the group conversation; a score for each message representing the relativity to the topics, length of the messages, the relationship between the messages and the topic calculated based on the Jaccard similarity; and the average word embedding for the message.

Another related work [4] is the multimodal real time and retrospective tool developed at University of Northwestern called "chemical pods," which assists professors in understanding student behaviors under active learning environment. The professors are able to access the audio analysis collected by the tool on the mobile app or server. The dashboard contains the following features: direction of arrival (DoA), speech recognition, keyword detection, question detection, sentiment analysis, user interactivity and speaker diarization. The design is being further developed to improve the speech-to-text accuracy and to incorporate body pose estimation. Our initial project idea was inspired by chemical pods and the desire to extend it to the collaborative gaming environment.

The results and concerns mentioned in these papers are taken into consideration when we are designing the tool and the evaluation system behind it.

3. Problem Description and Limitations

The City Settlers game has been deployed and tested in real classrooms before we started to design this evaluation system. The initial evaluation form was a survey filled by participants after each test play and the feedbacks from the instructors who walk around the classroom during the game. Therefore, the data collected is one-sided, limited by impressions and not reflective of the real effect of the game design.

3.1 Bias on Game-based Learning

The number of researches on game-based learning is increasing in recent years, yet according to the summary of [5], there are no definitively positive or negative results reported for a game-based learning environment. Instead, the 67% of positive estimates hardly prove the positive effect of game-based learning environment, and the situation varies from case to case.

In order to evaluate our game model for City Settlers comprehensively, we intend to collect data from all scales of the game.

3.2 Difficulties on Assessing Game-based Learning Environment

It is difficult to cover all the related aspects and details in assessing the efficiency of the game-based learning environment. An educational game may have various effects on different groups according to age, sex, educational background, and learning goals. The most common evaluation forms are in surveys, assignments, or quizzes after the activity, where participants recall and reflect on what they have experienced. However, self-evaluation may not adequately identify the benefit to participants or the advantages and disadvantages of the design.

Participant performance under the game-based learning environment is what we should focus on for retrieving the most accurate and reflective data in order to assess the efficiency of design.

3.3 Limitations

This section addresses concerns involved in the design of the tool for collecting the data.

3.3.1 Speech Separation

Speech separation is one of the most vexing concerns. The target group for City Settlers is teenagers attending middle school. Therefore, the surroundings for text-collection are chaotic which makes speech separation nearly impossible since groups of kids may frequently talk at the same time. The recordings

for previous test runs require manual speech-to-text because existing speech-to-text tools can hardly report the content. For our design of the tool, clearer sound recordings and speech separations are necessary for extracting the main content of the conversations.

3.3.2 Instant Feedback

Instant feedback is another concern in the aspects of time management and real-time performance analysis.

Consider the following scenarios:

1. When participants have a question during the collaboration, they have to go ask the instructors or wait for them to pass by, which is time-wasting for both sides.
2. If a group of participants are doing something unrelated to the class, instructors often know nothing about it.
3. Participants can merely remember what has happened in class so they cannot give a real reflection of the learning environment.

Adding instant feedback to the design would enhance the quality of the data collection and make a more reflective use of it.

4. Methods & Design

The goal for the design is to create a better environment for the City Settlers game, or apply the research results to improve game-based learning environments. There are three main features in this design: real-time feedback during the game, comprehensive performance analysis report based on the data collected, and an interactive platform for the instructors to access the data.

4.1 Real-time Feedback

The data collection tool is built on a Respeaker USB Mic Array from the Seeed Studio, which can record the conversation for each group during the game in high resolution. All the audios are stored locally in a Raspberry pi and uploaded to the cloud once the pi receives the audio file. The speech-to-text process is done through the Microsoft Azure cognitive service, and all the audio and text data are stored in the cloud. The work flow is shown in Figure 1.

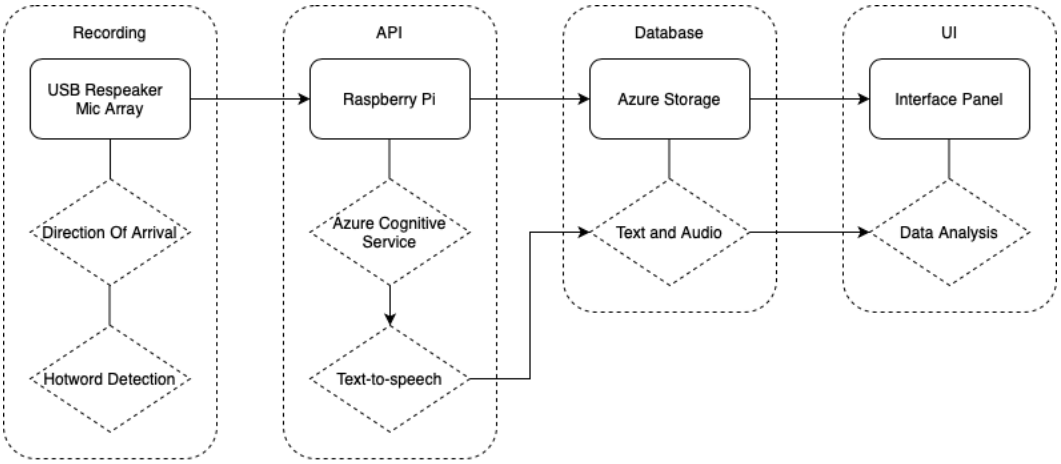

Figure 1 The Work Flow of the Tool

The Respeaker is set to detect if there is any incoming sound. The recording stops when a piece of conversation ends, and starts again when a new conversation begins. The audio input is stored in a newly created file and sent to the server right away for further analysis. In this way, the information that the instructors receive in the interface is up-to-date and they can keep track of recent topics in discussions and students' activities even without paying specific attention to one group.

I tried out several different methods for speech to text including the Baidu Deep Speech, Google Cloud speech-to-text and others. However, the accuracy for the speech-to-text is concerning since there are many translation mistakes on keywords that would affect further analysis. I decided to use the Microsoft

Azure Cognitvie Service since the accuracy is higher than the others in translating our audio files and it allows me to add keywords that may occur in the audio to increase the accuracy.

4.2 Comprehensive Performance Analysis

For data analysis, I use techniques from Natural Language Processing to analyze topics and sentiments for the contents that I require in text from the cloud. The main features we currently have are term frequencies, positive and negative discussions, questions, turn of talks, problem solving process and conclusion ideas such as the goal of the group, and some game logics that the participants are following during the game.

4.2.1 Topic Analysis

Since this tool is designed for the City Settlers game, we are able to summarize a specific set of words that frequently appear during the game. For instance, the words “copper”, “clay”, “food”, and “lumber” stand for resources, “clay mine”, “copper mine”, “food farm”, “fishing”, “food hunting”, and “lumber camp” are building names. Based on these keywords, we can filter the related information from all the data that we collect.

When I was choosing the algorithm for topic analysis, I decided to use Term Frequency and Inverse Document Frequency (TF-IDF) over Bag of Words based on the performances of the collected data from test runs. Since the data sizes are usually long but vary in contents in our case, the TF-IDF model contains information on the more important words and the less important ones as well compared to Bag of Words which creates a set of vectors containing the count of word occurrences in the document. Using TF-IDF can cover more contents and avoid losing data when analyzing off-topic behaviors of participants.

The text is first tokenized to singular word; then lemmatized to transform all the words with same meaning but in different time form into the same form; next, stemming is applied to the text for reduce plural words into single words; finally, the TF-IDF model is used on the text, in which term frequency counts for the appearing rate of the keywords in each document, and inverse document frequency will penalize for contents that are not related like the, a, an.

4.2.2 Sentiment Analysis

Analyzing positivity and negativity in dialogues requires sentiment analysis for the meaning of singular words and meaning of words putting together in sentences. Emotional detection in our

case is limited to symbols such as question marks and exclamation due to the lack of expression in audios. Therefore, we will focus more on text sentiment analysis. According to report [6], different deep learning models are the ideal approaches in the case.

In our design, the sentimental analysis is mainly used to detect the positivity and negativity of the content. If the conversation is related to the game play, it will be determined as positive content; if the conversation is off-topic, that the conversation is something not related to the game, it will be determined as negative content. However, the negativity also applies to the attitude of the participants. For instance, if the conversation is a fighting argument between participants, it will be determined as negative content.

After similar data processing steps as topic analysis, which removes punctuations, converts terms to the same form and case, and cleans stop words, I built our training set and testing set based on a bag of words model created based on the data. I implemented three machine learning models: Logistic Regression, Naïve Bayes and SVM, and decide to use the SVM model which has the best accuracy score.

4.2.3 Turn of Talks

Turn of talks are designed to see how frequent each participant in the group is talking and weigh the equitable distribution discourse. Since our tool is designed for the entire group instead of one participant, we intend to track down the performances of individuals during the game. Some participants may not actively engage in the game due to lack of interests or being attracted by other things during the game, which is not the intention as the collaborative game-based learning environments are designed for.

The tool not only involves the data from text but also the Direction of Arrival (DoA) data collected by the USB Respeaker Mic Array with the audio file. We intend to read the position data which includes direction and distance of the incoming sound to diagnose the identity of the participant. When the Respeaker is collecting audio data, it also marks the direction and the distance of the incoming sound. However, the audio data is separated from the DoA data. Therefore, we record the timeline for the DoA data and combine it with the audio data. Each piece of text belongs to a single speaker.

4.2.4 Problem Solving

Problem solving is one of the education goals of City Settlers, to teach participants how to manage lack of resources, disagreements within groups, and improve their communication skills. The progress for the problem solving is a significant part of participants' gain from the game. Tracking the cause, process towards solving the problem, and the solution in the end can build study cases for instructors to understand how participants comprehend the elements in the game and how they interact with the game.

There are no existing algorithm that can analyze problem solving in our specific design. I am able to diagnoise when the participants are facing questions and what the questions are. However, analyzing the solving process and solutions still requires manual interference. This will be one of the future goals of the design.

4.3 Interactive Access to Data Collection

In this section, we will explain the data visualization techniques we use in the user interface in order to help instructors better understand the collected data and improve the collaborative game-based learning environment. I intend to design a user-friendly platform for easy handling and convenient access.

4.3.1 Panel View

In the panel view of the interactive platform (see Figure 2), each group is listed along with the topic they are discussing, number of positive, and negative conversations, number of questions, percentage of active time, and a performance score determined by comprehensive factors.

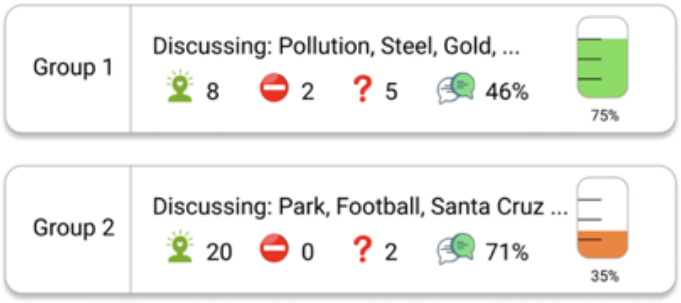

Figure 2 The Panel View of the User Interface

The panel view intends to give instructors an overview for the main activities of each group and how well they are doing. The performance score is determined by the following factors: how relatable are the topics the group is discussing with the game, the positivity and negativity of the discussions, and the active time of the group. If the performance score is below 50%, the platform will send an alert to the instructor.

4.3.2 Hotwords View

In each group analysis page, there is more detailed information with data and visualizations. Hotwords are displayed in a view of the line chart which emphasizes the appearance of words relating to the timeline. Instructors are able to view all the hotwords and its frequency in the game. Hotwords can help instructors to track the topics that the group has discussed and understand their interests. The view of hotwords is shown in Figure 3.

Figure 3 The Hotwords Detailed View

4.3.3 Active Time View

For active time, the visualization is in a radar graph (shown in Figure 4) displaying the participation rate for each individual member in the group. The active time for each participant is analyzed from the DoA report which identifies each individual by their orientation and distance. Speech separation is practiced with the help of DoA so that when individuals speak at the same time, their voices are labeled by their sound DoA.

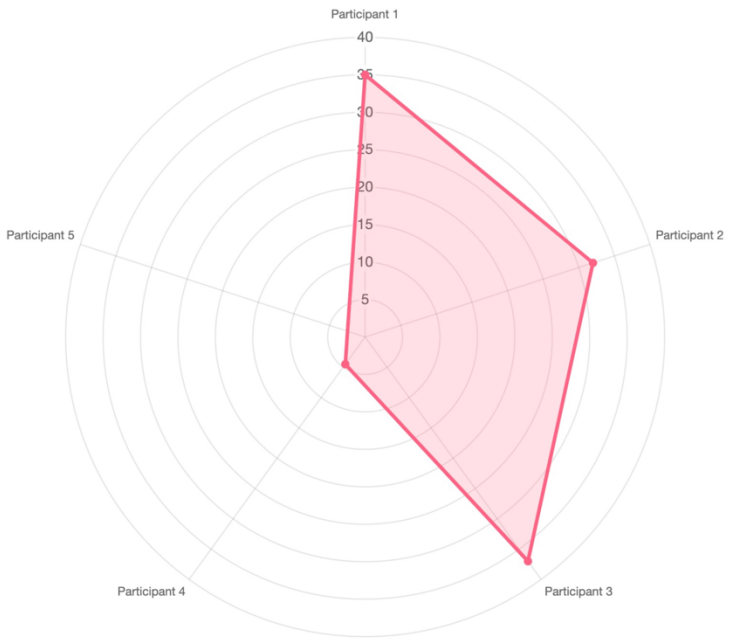

Figure 4 The Active Time Detailed View

When I initially designed the tool on Respeaker Core v2.0, which is another high resolution recorder designed by the Seeed Studio, I was not able to separate different voices and locate to each speaker. Voice separation is still a challenge especially for an environment that includes multiple voices and is chaotic overall. That is why I decide to use physical position detection to regnotize the identity of the speaker so it will be more accurate and easier than voice line separation. I also decide to change the recorder for the USB Respeaker Mic Array contains all the features as a recorder with convenient accesss to local device, clearer audio qualities, and as a complete design with an enclosure. So it can be carried around and used in classrooms without excessive worries on damaging the device.

5. Conclusion

We present this application to improve collaborative game-based learning environment for making it more controllable for the instructors and more responsible for the participants. With this application, participants and instructors are able to track performance in details for what they have done and what they should do. Unfortunately, our data is limited to the test runs with speech-to-text data translated manually, yet we are looking forward to utilize this tool and help the data collection in the future. The collaborative game-based environment needs further development and experiments to explore its real strength.

References

[1] R. Coley, J. Cradler and P. Engel, "Computers and Classrooms: The Status of Technology in U.S. Schools," ETS Policy Information Center, Princeton, New Jersey, 1997.

[2] J. Velvin, M. G. Gundersen Engeset, O. H. Graven and L. MacKinnon, "Using Games for Learning to Improve Students Performance in Higher Education," 2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE), Wollongong, NSW, 2018, pp. 674-678, doi: 10.1109/TALE.2018.8615323.

[3] D., Carpenter, A., Emerson, B., Mott, A., Saleh, K., Glazewski, C., Hmelo-Silver and J., Lester, 2020, "Detecting Off-Task Behavior from Student Dialogue in Game-Based Collaborative Learning," 21st International Conference on Artificial Intelligence in Education, [online] Springer, Volume 12163 LNAI, 2020, pp. 55-66.

[4] K., Anderson, T., Dubiel, K., Tanaka, and Worsley, M., 2019, "Chemistry Pods: A Multimodal Real Time and Retrospective Tool for the Classroom," 2019 International Conference on Multimodal Interaction (ICMI '19), October 14–18, 2019, Suzhou, China, ACM, New York, NY, USA, pp. 506-507, https://doi.org/10.1145/3340555. 3358662

[5] U., Tokac, E., Novak and C. G. Thompson, 2019, "Effects of Game-based Learning on Students' Mathematics Achievement: A Meta-analysis," Journal of Computer Assisted Learning, 35(3), pp. 407–420. doi: 10.1111/jcal.12347.

[6] K. K. Mohbey, "Sentiment Analysis for Product Rating Using a Deep Learning Approach," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), 2021, pp. 121-126, doi: 10.1109/ICAIS50930.2021.9395802.